

Serial Number: 09/445,803

**ENTERED**

1711  
#3

- ☒ Changed a file from non-ASCII to ASCII
- ☒ Changed the margins in cases where the sequence text was "wrapped" down to the next line.
- ☐ Edited a format error in the Current Application Data section, specifically: \_\_\_\_\_
- ☐ Edited the Current Application Data section with the actual current number. The number inputted by the applicant was ☐ the prior application data; or ☐ other \_\_\_\_\_
- ☐ Added the mandatory heading and subheadings for "Current Application Data".
- ☐ Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.
- ☐ Changed the spelling of a mandatory field (the headings or subheadings), specifically: \_\_\_\_\_
- ☐ Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were: \_\_\_\_\_
- ☐ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited: \_\_\_\_\_
- ☐ Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.
- ☐ Inserted colons after headings/subheadings. Headings edited included: \_\_\_\_\_
- ☐ Deleted extra, invalid, headings used by an applicant, specifically: \_\_\_\_\_
- ☐ Deleted: ☐ non-ASCII "garbage" at the beginning/end of files; ☐ secretary initials/filename at end of file; ☐ page numbers throughout text; ☐ other invalid text, such as \_\_\_\_\_
- ☐ Inserted mandatory headings, specifically: \_\_\_\_\_
- ☐ Corrected an obvious error in the response, specifically: \_\_\_\_\_
- ☐ Edited identifiers where upper case is used but lower case is required, or vice versa.
- ☐ Corrected an error in the Number of Sequences field, specifically: \_\_\_\_\_
- ☐ A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.
- ☐ Deleted *ending* stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (error due to a PatentIn bug). Sequences corrected: \_\_\_\_\_
- ☐ Other: \_\_\_\_\_

RECEIVED  
JUN 15 2001  
TO 1100 MAIL ROOM

\*Examiner: ~~The above~~ corrections must be communicated to the applicant in the first Office Action. DO NOT send a copy of this form.

RAW SEQUENCE LISTING                      DATE: 05/22/2000  
 PATENT APPLICATION: US/09/445,803        TIME: 18:53:32

Input Set : A:\Pto.amc  
 Output Set: N:\CRF3\05222000\I445803.raw

```

4 (1) GENERAL INFORMATION:
6   (i) APPLICANT: Adminis. of Tulane Educational, Fund
7       Philipp, Mario T.
9   (ii) TITLE OF INVENTION: Surface Antigens and Proteins Useful in
10      Compositions for the Diagnosis and Prevention of Lyme
11      Disease
13  (iii) NUMBER OF SEQUENCES: 14
15  (iv) CORRESPONDENCE ADDRESS:
16      (A) ADDRESSEE: Howson and Howson
17      (B) STREET: Spring House Corporate Cntr., P.O. Box 457
18      (C) CITY: Spring House
19      (D) STATE: Pennsylvania
20      (E) COUNTRY: USA
21      (F) ZIP: 19477
23  (v) COMPUTER READABLE FORM:
24      (A) MEDIUM TYPE: Floppy disk
25      (B) COMPUTER: IBM PC compatible
26      (C) OPERATING SYSTEM: PC-DOS/MS-DOS
27      (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
C--> 29  (vi) CURRENT APPLICATION DATA:
C--> 30      (A) APPLICATION NUMBER: US/09/445,803
C--> 31      (B) FILING DATE: 13-Dec-1999
32      (C) CLASSIFICATION:
34  (vii) PRIOR APPLICATION DATA:
35      (A) APPLICATION NUMBER: US 60/051,271
36      (B) FILING DATE: 30-JUN-1997
38  (viii) ATTORNEY/AGENT INFORMATION:
39      (A) NAME: Bak, Mary E.
40      (B) REGISTRATION NUMBER: 31,215
41      (C) REFERENCE/DOCKET NUMBER: TUL2APCT
43  (ix) TELECOMMUNICATION INFORMATION:
44      (A) TELEPHONE: 215-540-9200
45      (B) TELEFAX: 215-540-5818
48 (2) INFORMATION FOR SEQ ID NO: 1:
50   (i) SEQUENCE CHARACTERISTICS:
51      (A) LENGTH: 1047 base pairs
52      (B) TYPE: nucleic acid
53      (C) STRANDEDNESS: double
54      (D) TOPOLOGY: unknown
56   (ii) MOLECULE TYPE: cDNA
59   (ix) FEATURE:
60      (A) NAME/KEY: CDS
61      (B) LOCATION: 1..1047
64   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
66 AAG AAT AAT GAT CAT GAT AAT CAT AAG GGG ACT GTT AAG AAT GCT GTT 48
67 Lys Asn Asn Asp His Asp Asn His Lys Gly Thr Val Lys Asn Ala Val
68   1               5               10               15

```

## RAW SEQUENCE LISTING

DATE: 05/22/2000

PATENT APPLICATION: US/09/445,803

TIME: 18:53:32

Input Set : A:\Pto.amc

Output Set: N:\CRF3\05222000\I445803.raw

```

70 GAT ATG GCA AAG GCC GCT GAG GAA GCT GCA AGT GCT GCA AGT GCT GCT 96
71 Asp Met Ala Lys Ala Ala Glu Glu Ala Ala Ser Ala Ala Ser Ala Ala
72          20          25          30
74 ACT GGT AAT GCA GCG ATT GGG GAT GTT GTT AAG AAT AGT GGG GCA GCA 144
75 Thr Gly Asn Ala Ala Ile Gly Asp Val Val Lys Asn Ser Gly Ala Ala
76          35          40          45
78 GCA AAA GGT GGT GAG GCG GCG AGT GTT AAT GGG ATT GCT AAG GGG ATA 192
79 Ala Lys Gly Gly Glu Ala Ala Ser Val Asn Gly Ile Ala Lys Gly Ile
80          50          55          60
82 AAG GGG ATT GTT GAT GCT GCT GGA AAG GCT GAT GCG AAG GAA GGG AAG 240
83 Lys Gly Ile Val Asp Ala Ala Gly Lys Ala Asp Ala Lys Glu Gly Lys
84 65          70          75          80
86 TTG GAT GCT ACT GGT GCT GAG GGT ACG ACT AAC GTG AAT GCT GGG AAG 288
87 Leu Asp Ala Thr Gly Ala Glu Gly Thr Thr Asn Val Asn Ala Gly Lys
88          85          90          95
90 TTG TTT GTG AAG AGG GCG GCT GAT GAT GGT GGT GAT GCA GAT GAT GCT 336
91 Leu Phe Val Lys Arg Ala Ala Asp Asp Gly Gly Asp Ala Asp Asp Ala
92          100          105          110
94 GGG AAG GCT GCT GCT GCG GTT GCT GCA AGT GCT GCT ACT GGT AAT GCA 384
95 Gly Lys Ala Ala Ala Ala Val Ala Ala Ser Ala Ala Thr Gly Asn Ala
96          115          120          125
98 GCG ATT GGA GAT GTT GTT AAT GGT GAT GTG GCA AAA GCA AAA GGT GGT 432
99 Ala Ile Gly Asp Val Val Asn Gly Asp Val Ala Lys Ala Lys Gly Gly
100          130          135          140
102 GAT GCG GCG AGT GTT AAT GGG ATT GCT AAG GGT ATA AAG GGG ATT GTT 480
103 Asp Ala Ala Ser Val Asn Gly Ile Ala Lys Gly Ile Lys Gly Ile Val
104 145          150          155          160
106 GAT GCT GCT GAG AAG GCT GAT GCG AAG GAA GGG AAG TTG AAT GCT GCT 528
107 Asp Ala Ala Glu Lys Ala Asp Ala Lys Glu Gly Lys Leu Asn Ala Ala
108          165          170          175
110 GGT GCT GAG GGT ACG ACT AAC GCG GAT GCT GGG AAG TTG TTT GTG AAG 576
111 Gly Ala Glu Gly Thr Thr Asn Ala Asp Ala Gly Lys Leu Phe Val Lys
112          180          185          190
114 AAT GCT GGT AAT GTG GGT GGT GAA GCA GGT GAT GCT GGG AAG GCT GCT 624
115 Asn Ala Gly Asn Val Gly Gly Glu Ala Gly Asp Ala Gly Lys Ala Ala
116          195          200          205
118 GCT GCG GTT GCT GCT GTT AGT GGG GAG CAG ATA TTA AAA GCG ATT GTT 672
119 Ala Ala Val Ala Ala Val Ser Gly Glu Gln Ile Leu Lys Ala Ile Val
120          210          215          220
122 CAT GCT GCT AAG GAT GGT GGT GAG AAG CAG GGT AAG AAG GCT GCG GAT 720
123 His Ala Ala Lys Asp Gly Gly Glu Lys Gln Gly Lys Lys Ala Ala Asp
124 225          230          235          240
126 CGT ACA AAT CCC ATT GAC GCG GCT ATT GGG GGT GCG GGT GAT AAT GAT 768
127 Arg Thr Asn Pro Ile Asp Ala Ala Ile Gly Gly Ala Gly Asp Asn Asp
128          245          250          255
130 GCT GCT GCG GCG TTT GCT ACT ATG AAG AAG GAT GAT CAG ATT GCT GCT 816
131 Ala Ala Ala Ala Phe Ala Thr Met Lys Lys Asp Asp Gln Ile Ala Ala
132          260          265          270
134 GCT ATG GTT CTG AGG GGA ATG GCT AAG GAT GGG CAA TTT GCT TTG AAG 864

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/445,803

DATE: 05/22/2000

TIME: 18:53:32

Input Set : A:\Pto.amc

Output Set: N:\CRF3\05222000\I445803.raw

```

135 Ala Met Val Leu Arg Gly Met Ala Lys Asp Gly Gln Phe Ala Leu Lys
136      275      280      285
138 GAT GCT GCT GCT GCT CAT GAA GGG ACT GTT AAG AAT GCT GTT GAT ATA 912
139 Asp Ala Ala Ala Ala His Glu Gly Thr Val Lys Asn Ala Val Asp Ile
140      290      295      300
142 ATA AAG GCT GCT GCG GAA GCT GCA AGT GCT GCA AGT GCT GCT ACT GGT 960
143 Ile Lys Ala Ala Ala Glu Ala Ala Ser Ala Ala Ser Ala Ala Thr Gly
144 305      310      315      320
146 AGT GCA GCA ATT GGG GAT GTT GTT AAT GGT AAT GGA GCA ACA GCA AAA 1008
147 Ser Ala Ala Ile Gly Asp Val Val Asn Gly Asn Gly Ala Thr Ala Lys
148      325      330      335
150 GGT GGT GAT GCG AAG AGT GTT AAT GGC ATT GCT AAG GGA      1047
151 Gly Gly Asp Ala Lys Ser Val Asn Gly Ile Ala Lys Gly
152      340      345
155 (2) INFORMATION FOR SEQ ID NO: 2:
157   (i) SEQUENCE CHARACTERISTICS:
158       (A) LENGTH: 349 amino acids
159       (B) TYPE: amino acid
160       (D) TOPOLOGY: linear
162   (ii) MOLECULE TYPE: protein
164   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
166 Lys Asn Asn Asp His Asp Asn His Lys Gly Thr Val Lys Asn Ala Val
167 1      5      10      15
169 Asp Met Ala Lys Ala Ala Glu Glu Ala Ala Ser Ala Ala Ser Ala Ala
170      20      25      30
172 Thr Gly Asn Ala Ala Ile Gly Asp Val Val Lys Asn Ser Gly Ala Ala
173      35      40      45
175 Ala Lys Gly Gly Glu Ala Ala Ser Val Asn Gly Ile Ala Lys Gly Ile
176      50      55      60
178 Lys Gly Ile Val Asp Ala Ala Gly Lys Ala Asp Ala Lys Glu Gly Lys
179 65      70      75      80
181 Leu Asp Ala Thr Gly Ala Glu Gly Thr Thr Asn Val Asn Ala Gly Lys
182      85      90      95
184 Leu Phe Val Lys Arg Ala Ala Asp Asp Gly Gly Asp Ala Asp Asp Ala
185      100      105      110
187 Gly Lys Ala Ala Ala Val Ala Ala Ser Ala Ala Thr Gly Asn Ala
188      115      120      125
190 Ala Ile Gly Asp Val Val Asn Gly Asp Val Ala Lys Ala Lys Gly Gly
191      130      135      140
193 Asp Ala Ala Ser Val Asn Gly Ile Ala Lys Gly Ile Lys Gly Ile Val
194 145      150      155      160
196 Asp Ala Ala Glu Lys Ala Asp Ala Lys Glu Gly Lys Leu Asn Ala Ala
197      165      170      175
199 Gly Ala Glu Gly Thr Thr Asn Ala Asp Ala Gly Lys Leu Phe Val Lys
200      180      185      190
202 Asn Ala Gly Asn Val Gly Gly Glu Ala Gly Asp Ala Gly Lys Ala Ala
203      195      200      205
205 Ala Ala Val Ala Ala Val Ser Gly Glu Gln Ile Leu Lys Ala Ile Val
206      210      215      220

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/445,803

DATE: 05/22/2000

TIME: 18:53:32

Input Set : A:\Pto.amc

Output Set: N:\CRF3\05222000\I445803.raw

```

208 His Ala Ala Lys Asp Gly Gly Glu Lys Gln Gly Lys Lys Ala Ala Asp
209 225          230          235          240
211 Arg Thr Asn Pro Ile Asp Ala Ala Ile Gly Gly Ala Gly Asp Asn Asp
212          245          250          255
214 Ala Ala Ala Ala Phe Ala Thr Met Lys Lys Asp Asp Gln Ile Ala Ala
215          260          265          270
217 Ala Met Val Leu Arg Gly Met Ala Lys Asp Gly Gln Phe Ala Leu Lys
218          275          280          285
220 Asp Ala Ala Ala Ala His Glu Gly Thr Val Lys Asn Ala Val Asp Ile
221          290          295          300
223 Ile Lys Ala Ala Ala Glu Ala Ala Ser Ala Ala Ser Ala Ala Thr Gly
224 305          310          315          320
226 Ser Ala Ala Ile Gly Asp Val Val Asn Gly Asn Gly Ala Thr Ala Lys
227          325          330          335
229 Gly Gly Asp Ala Lys Ser Val Asn Gly Ile Ala Lys Gly
230          340          345
232 (2) INFORMATION FOR SEQ ID NO: 3:
234 (i) SEQUENCE CHARACTERISTICS:
235 (A) LENGTH: 283 base pairs
236 (B) TYPE: nucleic acid
237 (C) STRANDEDNESS: double
238 (D) TOPOLOGY: unknown
240 (ii) MOLECULE TYPE: DNA (genomic)
245 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
247 GCCGCTGGAT GGTGGTGAGA AGCAGGGTAA GAAGGCTGCG GATCGTACAA ATCCCATTTGA 60
249 CGCGGCTATT GGGGGTGCGG GTGATAATGA TGCTGCTGCG GCGTTTGCTA CTATGAAGAA 120
251 GGATGATCAG ATTGCTGCTG CTATGGTTCT GAGGGGAATG GCTAAGGATG GGCAATTGTC 180
253 TTTGAAGGAT GCTGCTGCTG CTCATGAAGG GACTGTTAAG AATGCTGTTG ATATAATAAA 240
255 GGCTGCTGCG GAAGCTGCAA GTGCTGCAAG TGCTGCTACT GGT 283
257 (2) INFORMATION FOR SEQ ID NO: 4:
259 (i) SEQUENCE CHARACTERISTICS:
260 (A) LENGTH: 233 base pairs
261 (B) TYPE: nucleic acid
262 (C) STRANDEDNESS: double
263 (D) TOPOLOGY: unknown
265 (ii) MOLECULE TYPE: DNA (genomic)
270 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
272 TTTATTATAT CAACAGATTC TTAACAGTCC CTTTCATGAGC AGCAGCAGCA TCCTTCAAAG 60
274 CAAATTGCCC ATCCTTAGCC ATTCCCCTCA GAACCATAGC AGCAGCAATC TGATCATCCT 120
276 TCTTCATAGT AGCAAACGCC GCAGCAGCAT CATTATCACC GCACCCCCCA ATAGCCGCGT 180
278 CAATCGGATT TGTACGATCC GCAGCCTTCT TACCCTGCTT CTCACCACCA TCC 233
280 (2) INFORMATION FOR SEQ ID NO: 5:
282 (i) SEQUENCE CHARACTERISTICS:
283 (A) LENGTH: 194 base pairs
284 (B) TYPE: nucleic acid
285 (C) STRANDEDNESS: double
286 (D) TOPOLOGY: unknown
288 (ii) MOLECULE TYPE: DNA (genomic)
293 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/445,803  
 DATE: 05/22/2000  
 TIME: 18:53:32

Input Set : A:\Pto.amc  
 Output Set : N:\CRF3\05222000\I445803.raw

```

295 CCGTGCAAGC TGGGTTGAAG AAGGTTGGGG ATGTTGTTAA GAATAGTGAG GCAAAAGATG 60
297 GTGATGCCGC GAGTGTAAAT GGGATTGCTA AGGGGATAAA GGGGATTGTT GATGCTGCTG 120
299 AGAAGGCTGA TCGGAAGGAA GGGAAGTTGG TATGTGGCTG GTGCTGCTGG TGAAACTAAC 180
301 AAGGAAGCGG CCGC 194
303 (2) INFORMATION FOR SEQ ID NO: 6:
305 (i) SEQUENCE CHARACTERISTICS:
306 (A) LENGTH: 369 base pairs
307 (B) TYPE: nucleic acid
308 (C) STRANDEDNESS: double
309 (D) TOPOLOGY: unknown
311 (ii) MOLECULE TYPE: DNA (genomic)
316 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
318 GCGGCCGCTT GAGGAAGCTG CAAGTGCTGC AAGTGCTGCT ACTGGTAATG CAGCGATTGG 60
320 GGATGTTGTT AAGAAATAGT GGGCAGCAGC AAAAGGTTGG GAGGCGGCCA GTGTTAATGG 120
322 GATTGCTAAG GGGATAAAGG GGATTGTTGA TGCTGCTGGA AAGGCTGATG CGAAGGAAGG 180
324 GAAGTTGGAT GCTACTGGTG CTGAGGGTAC GACTAACGTG AATGCTGGGA AGTGTGTTGT 240
326 GAAGAGGGCG GCTGATGATG GTGGTGTATC AGATGATGCT GGGAAGGCTG CTGCTGCGGT 300
328 TGCTGCAAGT GCTGCTACTG GTAATGCAGC GATTGGAGAT GTTGTTAATG GTGATGTGGC 360
330 AAAACAAA 369
332 (2) INFORMATION FOR SEQ ID NO: 7:
334 (i) SEQUENCE CHARACTERISTICS:
335 (A) LENGTH: 142 base pairs
336 (B) TYPE: nucleic acid
337 (C) STRANDEDNESS: double
338 (D) TOPOLOGY: unknown
340 (ii) MOLECULE TYPE: DNA (genomic)
345 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
347 AAGGATGGTG ATGATAAGCA GGGTAAGAAG GCTGAGGATG CTACAAATCC GATTGACGCG 60
349 GCTATGCGGG GTGCAGGTGC GGGTGCTAAT GCTGCTGCGG CGTTTAATAA TATGAAGAAG 120
351 GATGATCAGA TTGAGCGGCC GC 142
353 (2) INFORMATION FOR SEQ ID NO: 8:
355 (i) SEQUENCE CHARACTERISTICS:
356 (A) LENGTH: 210 base pairs
357 (B) TYPE: nucleic acid
358 (C) STRANDEDNESS: double
359 (D) TOPOLOGY: unknown
361 (ii) MOLECULE TYPE: DNA (genomic)
366 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
368 GGTGAACTA ACAAGGATGC TGGGAAGTTG TTTGTGAAGA AGAATGGTGA TGATGGTGGT 60
370 GATGCAGGTG ATGCTGGGAA GGCTGCTGCT GCGGTTGCTG CTGTTAGTGG GGAGCAGATA 120
372 TTAAAGCGA TTGTTGATGC TGCTAAAGAT GGTGATAAGA CGGGGGTTAC TGATGTAAAG 180
374 GATGCTACAA ATCCGATTGA CGCGGCTATT 210
376 (2) INFORMATION FOR SEQ ID NO: 9:
378 (i) SEQUENCE CHARACTERISTICS:
379 (A) LENGTH: 236 base pairs
380 (B) TYPE: nucleic acid
381 (C) STRANDEDNESS: double
382 (D) TOPOLOGY: unknown
384 (ii) MOLECULE TYPE: DNA (genomic)

```

1711

RAW SEQUENCE LISTING                      DATE: 05/12/2000  
 PATENT APPLICATION:    US/09/445,803              TIME: 11:04:23

Input Set : A:\sequence.lst.txt  
 Output Set: N:\CRF3\05122000\I445803.raw

4 (1) GENERAL INFORMATION:  
 6    (i) APPLICANT: Adminis. of Tulane Educational, Fund  
 7                      Philipp, Mario T.  
 9    (ii) TITLE OF INVENTION: Surface Antigens and Proteins Useful in  
 10                                      Compositions for the Diagnosis and Prevention of Lyme  
 11                                      Disease  
 13    (iii) NUMBER OF SEQUENCES: 14  
 15    (iv) CORRESPONDENCE ADDRESS: .  
 16                      (A) ADDRESSEE: Howson and Howson  
 17                      (B) STREET: Spring House Corporate Cntr., P.O. Box 457  
 18                      (C) CITY: Spring House  
 19                      (D) STATE: Pennsylvania  
 20                      (E) COUNTRY: USA  
 21                      (F) ZIP: 19477  
 23    (v) COMPUTER READABLE FORM:  
 24                      (A) MEDIUM TYPE: Floppy disk  
 25                      (B) COMPUTER: IBM PC compatible  
 26                      (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
 27                      (D) SOFTWARE: PatentIn Release #1.0, Version #1.30  
 C--> 29    (vi) CURRENT APPLICATION DATA:  
 C--> 30                      (A) APPLICATION NUMBER: US/09/445,803  
 C--> 31                      (B) FILING DATE: 13-Dec-1999  
 32                      (C) CLASSIFICATION:  
 34    (vii) PRIOR APPLICATION DATA:  
 35                      (A) APPLICATION NUMBER: US 60/051,271  
 36                      (B) FILING DATE: 30-JUN-1997  
 38    (viii) ATTORNEY/AGENT INFORMATION:  
 39                      (A) NAME: Bak, Mary E.  
 40                      (B) REGISTRATION NUMBER: 31,215  
 41                      (C) REFERENCE/DOCKET NUMBER: TUL2APCT  
 43    (ix) TELECOMMUNICATION INFORMATION:  
 44                      (A) TELEPHONE: 215-540-9200  
 45                      (B) TELEFAX: 215-540-5818

**Does Not Comply  
 Corrected Diskette Needed**

#### ERRORRED SEQUENCES

48 (2) INFORMATION FOR SEQ ID NO: 1:  
 50    (i) SEQUENCE CHARACTERISTICS:  
 51                      (A) LENGTH: 1047 base pairs  
 52                      (B) TYPE: nucleic acid  
 53                      (C) STRANDEDNESS: double  
 54                      (D) TOPOLOGY: unknown  
 56    (ii) MOLECULE TYPE: cDNA  
 59    (ix) FEATURE:  
 60                      (A) NAME/KEY: CDS  
 61                      (B) LOCATION: 1..1047

## RAW SEQUENCE LISTING

DATE: 05/12/2000

PATENT APPLICATION: US/09/445,803

TIME: 11:04:23

Input Set : A:\sequence.lst.txt

Output Set: N:\CRF3\05122000\I445803.raw

64 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

W--> 66 AAG AAT AAT GAT CAT GAT AAT CAT AAG GGG ACT GTT AAG AAT GCT GTT

W--> 67 48

68 Lys Asn Asn Asp His Asp Asn His Lys Gly Thr Val Lys Asn Ala Val

W--> 69 1 5 10 15

W--> 71 GAT ATG GCA AAG GCC GCT GAG GAA GCT GCA AGT GCT GCA AGT GCT GCT

W--> 72 96

73 Asp Met Ala Lys Ala Ala Glu Glu Ala Ala Ser Ala Ala Ser Ala Ala

W--> 74 20 25 30

W--> 76 ACT GGT AAT GCA GCG ATT GGG GAT GTT GTT AAG AAT AGT GGG GCA GCA

W--> 77 144

78 Thr Gly Asn Ala Ala Ile Gly Asp Val Val Lys Asn Ser Gly Ala Ala

W--> 79 35 40 45

W--> 81 GCA AAA GGT GGT GAG GCG GCG AGT GTT AAT GGG ATT GCT AAG GGG ATA

W--> 82 192

83 Ala Lys Gly Gly Glu Ala Ala Ser Val Asn Gly Ile Ala Lys Gly Ile

W--> 84 50 55 60

W--> 86 AAG GGG ATT GTT GAT GCT GCT GGA AAG GCT GAT GCG AAG GAA GGG AAG

W--> 87 240

88 Lys Gly Ile Val Asp Ala Ala Gly Lys Ala Asp Ala Lys Glu Gly Lys

W--> 89 65 70 75 80

W--> 91 TTG GAT GCT ACT GGT GCT GAG GGT ACG ACT AAC GTG AAT GCT GGG AAG

W--> 92 288

93 Leu Asp Ala Thr Gly Ala Glu Gly Thr Thr Asn Val Asn Ala Gly Lys

W--> 94 85 90 95

W--> 96 TTG TTT GTG AAG AGG GCG GCT GAT GAT GGT GGT GAT GCA GAT GAT GCT

W--> 97 336

98 Leu Phe Val Lys Arg Ala Ala Asp Asp Gly Gly Asp Ala Asp Asp Ala

W--> 99 100 105 110

W--> 101 GGG AAG GCT GCT GCT GCG GTT GCT GCA AGT GCT GCT ACT GGT AAT GCA

W--> 102 384

103 Gly Lys Ala Ala Ala Val Ala Ala Ser Ala Ala Thr Gly Asn Ala

W--> 104 115 120 125

W--> 106 GCG ATT GGA GAT GTT GTT AAT GGT GAT GTG GCA AAA GCA AAA GGT GGT

W--> 107 432

108 Ala Ile Gly Asp Val Val Asn Gly Asp Val Ala Lys Ala Lys Gly Gly

W--> 109 130 135 140

W--> 111 GAT GCG GCG AGT GTT AAT GGG ATT GCT AAG GGT ATA AAG GGG ATT GTT

W--> 112 480

113 Asp Ala Ala Ser Val Asn Gly Ile Ala Lys Gly Ile Lys Gly Ile Val

W--> 114 145 150 155 160

W--> 116 GAT GCT GCT GAG AAG GCT GAT GCG AAG GAA GGG AAG TTG AAT GCT GCT

W--> 117 528

118 Asp Ala Ala Glu Lys Ala Asp Ala Lys Glu Gly Lys Leu Asn Ala Ala

W--> 119 165 170 175

W--> 121 GGT GCT GAG GGT ACG ACT AAC GCG GAT GCT GGG AAG TTG TTT GTG AAG

W--> 122 576

123 Gly Ala Glu Gly Thr Thr Asn Ala Asp Ala Gly Lys Leu Phe Val Lys

W--> 124 180 185 190

48

wrapped text

96

↓



## RAW SEQUENCE LISTING

DATE: 05/12/2000

PATENT APPLICATION: US/09/445,803

TIME: 11:04:23

Input Set : A:\sequence.lst.txt

Output Set: N:\CRF3\05122000\I445803.raw

```

W--> 126 AAT GCT GGT AAT GTG GGT GGT GAA GCA GGT GAT GCT GGG AAG GCT GCT
W--> 127 624
      128 Asn Ala Gly Asn Val Gly Gly Glu Ala Gly Asp Ala Gly Lys Ala Ala
W--> 129          195          200          205
W--> 131 GCT GCG GTT GCT GCT GTT AGT GGG GAG CAG ATA TTA AAA GCG ATT GTT
W--> 132 672
      133 Ala Ala Val Ala Ala Val Ser Gly Glu Gln Ile Leu Lys Ala Ile Val
W--> 134          210          215          220
W--> 136 CAT GCT GCT AAG GAT GGT GGT GAG AAG CAG GGT AAG AAG GCT GCG GAT
W--> 137 720
      138 His Ala Ala Lys Asp Gly Gly Glu Lys Gln Gly Lys Lys Ala Ala Asp
W--> 139 225          230          235          240
W--> 141 CGT ACA AAT CCC ATT GAC GCG GCT ATT GGG GGT GCG GGT GAT AAT GAT
W--> 142 768
      143 Arg Thr Asn Pro Ile Asp Ala Ala Ile Gly Gly Ala Gly Asp Asn Asp
W--> 144          245          250          255
W--> 146 GCT GCT GCG GCG TTT GCT ACT ATG AAG AAG GAT GAT CAG ATT GCT GCT
W--> 147 816
      148 Ala Ala Ala Ala Phe Ala Thr Met Lys Lys Asp Asp Gln Ile Ala Ala
W--> 149          260          265          270
W--> 151 GCT ATG GTT CTG AGG GGA ATG GCT AAG GAT GGG CAA TTT GCT TTG AAG
W--> 152 864
      153 Ala Met Val Leu Arg Gly Met Ala Lys Asp Gly Gln Phe Ala Leu Lys
W--> 154          275          280          285
W--> 156 GAT GCT GCT GCT GCT CAT GAA GGG ACT GTT AAG AAT GCT GTT GAT ATA
W--> 157 912
      158 Asp Ala Ala Ala Ala His Glu Gly Thr Val Lys Asn Ala Val Asp Ile
W--> 159          290          295          300
W--> 161 ATA AAG GCT GCT GCG GAA GCT GCA AGT GCT GCA AGT GCT GCT ACT GGT
W--> 162 960
      163 Ile Lys Ala Ala Ala Glu Ala Ala Ser Ala Ala Ser Ala Ala Thr Gly
W--> 164 305          310          315          320
W--> 166 AGT GCA GCA ATT GGG GAT GTT GTT AAT GGT AAT GGA GCA ACA GCA AAA
W--> 167 1008
      168 Ser Ala Ala Ile Gly Asp Val Val Asn Gly Asn Gly Ala Thr Ala Lys
W--> 169          325          330          335
W--> 171 GGT GGT GAT GCG AAG AGT GTT AAT GGC ATT GCT AAG GGA
W--> 172 1047
      173 Gly Gly Asp Ala Lys Ser Val Asn Gly Ile Ala Lys Gly
E--> 174          340          345
254 (2) INFORMATION FOR SEQ ID NO: 3:
256 (i) SEQUENCE CHARACTERISTICS:
257 (A) LENGTH: 283 base pairs
258 (B) TYPE: nucleic acid
259 (C) STRANDEDNESS: double
260 (D) TOPOLOGY: unknown
262 (ii) MOLECULE TYPE: DNA (genomic)
267 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
E--> 269 GCCGCTGGAT GGTGGTGAGA AGCAGGGTAA GAAGGCTGCG GATCGTACAA ATCCCATTGA

```

*Same*

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/445,803  
 DATE: 05/12/2000  
 TIME: 11:04:23

Input Set : A:\sequence.lst.txt  
 Output Set: N:\CRF3\05122000\I445803.raw

W--> 270 60  
 E--> 272 CGCGGCTATT GGGGGTGC GG GTGATAATGA TGCTGCTGCG GCGTTTGCTA CTATGAAGAA  
 W--> 273 120  
 E--> 275 GGATGATCAG ATTGCTGCTG CTATGTTCT GAGGGGAATG GCTAAGGATG GGCAATTGCG  
 W--> 276 180  
 E--> 278 TTTGAAGGAT GCTGCTGCTG CTCATGAAGG GACTGTTAAG AATGCTGTTG ATATAATAAA  
 W--> 279 240  
 E--> 281 GGCTGCTGCG GAAGCTGCAA GTGCTGCAAG TGCTGCTACT GGT  
 E--> 282 283  
 284 (2) INFORMATION FOR SEQ ID NO: 4:  
 286 (i) SEQUENCE CHARACTERISTICS:  
 287 (A) LENGTH: 233 base pairs  
 288 (B) TYPE: nucleic acid  
 289 (C) STRANDEDNESS: double  
 290 (D) TOPOLOGY: unknown  
 292 (ii) MOLECULE TYPE: DNA (genomic)  
 297 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
 E--> 299 TTTATTATAT CAACAGATTC TTAACAGTCC CTTCATGAGC AGCAGCAGCA TCCTTCAAAG  
 W--> 300 60  
 E--> 302 CAAATTGCCC ATCCTTAGCC ATTCCCTCA GAACCATAGC AGCAGCAATC TGATCATCCT  
 W--> 303 120  
 E--> 305 TCTTCATAGT AGCAAACGCC GCAGCAGCAT CATTATCACC CGCACCCCCA ATAGCCGCGT  
 W--> 306 180  
 E--> 308 CAATCGGATT TGTACGATCC GCAGCCTTCT TACCCTGCTT CTCACCACCA TCC  
 E--> 309 233  
 311 (2) INFORMATION FOR SEQ ID NO: 5:  
 313 (i) SEQUENCE CHARACTERISTICS:  
 314 (A) LENGTH: 194 base pairs  
 315 (B) TYPE: nucleic acid  
 316 (C) STRANDEDNESS: double  
 317 (D) TOPOLOGY: unknown  
 319 (ii) MOLECULE TYPE: DNA (genomic)  
 324 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
 E--> 326 CCGTGCAAGC TGGGTGAAG AAGGTGGGG ATGTTGTAA GAATAGTGAG GCAAAAGATG  
 W--> 327 60  
 E--> 329 GTGATGCGGC GAGTGTTAAT GGGATTGCTA AGGGGATAAA GGGGATTGTT GATGCTGCTG  
 W--> 330 120  
 E--> 332 AGAAGGCTGA TGCGAAGGAA GGAAGTTGG TATGTGGCTG GTGCTGCTGG TGAAACTAAC  
 W--> 333 180  
 E--> 335 AAGGAAGCGG CCG  
 E--> 336 194  
 338 (2) INFORMATION FOR SEQ ID NO: 6:  
 340 (i) SEQUENCE CHARACTERISTICS:  
 341 (A) LENGTH: 369 base pairs  
 342 (B) TYPE: nucleic acid  
 343 (C) STRANDEDNESS: double  
 344 (D) TOPOLOGY: unknown  
 346 (ii) MOLECULE TYPE: DNA (genomic)  
 351 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:


*same*

RAW SEQUENCE LISTING                      DATE: 05/12/2000  
PATENT APPLICATION: US/09/445,803              TIME: 11:04:23

Input Set : A:\sequence.lst.txt  
Output Set: N:\CRF3\05122000\I445803.raw

E--> 353 GCGGCCGCTT GAGGAAGCTG CAAGTGCTGC AAGTGCTGCT ACTGGTAATG CAGCGATTGG  
W--> 354 60  
E--> 356 GGATGTTGTT AAGAATAGTG GGCAGCAGC AAAAGGTGGT GAGGCGGCCA GTGTTAATGG  
W--> 357 120  
E--> 359 GATTGCTAAG GGGATAAAGG GGATTGTTGA TGCTGCTGGA AAGGCTGATG CGAAGGAAGG  
W--> 360 180  
E--> 362 GAAGTTGGAT GCTACTGGTG CTGAGGGTAC GACTAACGTG AATGCTGGGA AGTTGTTTGT  
W--> 363 240  
E--> 365 GAAGAGGGCG GCTGATGATG GTGGTGATGC AGATGATGCT GGGGAAGGCTG CTGCTGCGGT  
W--> 366 300  
E--> 368 TGCTGCAAGT GCTGCTACTG GTAATGCAGC GATTGGAGAT GTTGTTAATG GTGATGTGGC  
W--> 369 360  
E--> 371 AAAACAAAA  
E--> 372 369  
374 (2) INFORMATION FOR SEQ ID NO: 7:  
376 (i) SEQUENCE CHARACTERISTICS:  
377 (A) LENGTH: 142 base pairs  
378 (B) TYPE: nucleic acid  
379 (C) STRANDEDNESS: double  
380 (D) TOPOLOGY: unknown  
382 (ii) MOLECULE TYPE: DNA (genomic)  
387 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
E--> 389 AAGGATGGTG ATGATAAGCA GGGTAAGAAG GCTGAGGATG CTACAAATCC GATTGACGCG  
W--> 390 60  
E--> 392 GCTATTGGGG GTGCAGGTGC GGGTGCTAAT GCTGCTGCGG CGTTTAATAA TATGAAGAAG  
W--> 393 120  
E--> 395 GATGATCAGA TTGAGCGGCC GC  
E--> 396 142  
398 (2) INFORMATION FOR SEQ ID NO: 8:  
400 (i) SEQUENCE CHARACTERISTICS:  
401 (A) LENGTH: 210 base pairs  
402 (B) TYPE: nucleic acid  
403 (C) STRANDEDNESS: double  
404 (D) TOPOLOGY: unknown  
406 (ii) MOLECULE TYPE: DNA (genomic)  
411 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
E--> 413 GGTGAACTA ACAAGGATGC TGGAAGTTG TTTGTGAAGA AGAATGGTGA TGATGGTGGT  
W--> 414 60  
E--> 416 GATGCAGGTG ATGCTGGGAA GGCTGCTGCT GCGGTTGCTG CTGTTAGTGG GGAGCAGATA  
W--> 417 120  
E--> 419 TTTAAAGCGA TTGTTGATGC TGCTAAAGAT GGTGATAAGA CGGGGGTTAC TGATGTAAAG  
W--> 420 180  
E--> 422 GATGCTACAA ATCCGATTGA CGCGGCTATT  
E--> 423 210  
425 (2) INFORMATION FOR SEQ ID NO: 9:  
427 (i) SEQUENCE CHARACTERISTICS:  
428 (A) LENGTH: 236 base pairs  
429 (B) TYPE: nucleic acid  
430 (C) STRANDEDNESS: double

*same*



RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/445,803  
 DATE: 05/12/2000  
 TIME: 11:04:23

Input Set : A:\sequence.lst.txt  
 Output Set: N:\CRF3\05122000\I445803.raw

431 (D) TOPOLOGY: unknown  
 433 (ii) MOLECULE TYPE: DNA (genomic)  
 438 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:  
 E--> 440 TATATAATAA AGGCTGCTGC GAAGCTGCAA GTGCTGCAAG TGCTGCTACT GGTAGTGCAG  
 W--> 441 60  
 E--> 443 CAATTGGGGA TGTGTGTAAT GGTAATGGAG CAACAGCAAA AGGTGGTGAT GCGAAGTGTT  
 W--> 444 120  
 E--> 446 AATGGGATTG CTAAGGGGAT AAAGGGGATT GTTGATGCTG CTGAGAAGGC TGATGCGAAG  
 W--> 447 180  
 E--> 449 GAAGGGAAGT TGGATGTGGC TGGTGATGCT GGTGAAACTA ACAAGGAAGC GGCCGC  
 E--> 450 236  
 452 (2) INFORMATION FOR SEQ ID NO: 10:  
 454 (i) SEQUENCE CHARACTERISTICS:  
 455 (A) LENGTH: 199 base pairs  
 456 (B) TYPE: nucleic acid  
 457 (C) STRANDEDNESS: double  
 458 (D) TOPOLOGY: unknown  
 460 (ii) MOLECULE TYPE: DNA (genomic)  
 465 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
 E--> 467 ATGAGAGGAT CTCATCACCA TCACCATCAC ACGGATCCCC CGGGCTGCAG GAATTCGCGG  
 W--> 468 60  
 E--> 470 CCGCTGAAGG CTGATGCGAA GGAAGGGAAG TTGGATGTGG CTGGTGCTGC TGGTGAACAT  
 W--> 471 120  
 E--> 473 AACAGGATG CTGGGAAGTT GTTGTGAAG AAGAATAATG AGGGTGGTGA AGCAAATGAT  
 W--> 474 180  
 E--> 476 GCTGGGAAGG CTGCTGCTG  
 E--> 477 199  
 479 (2) INFORMATION FOR SEQ ID NO: 11:  
 481 (i) SEQUENCE CHARACTERISTICS:  
 482 (A) LENGTH: 272 base pairs  
 483 (B) TYPE: nucleic acid  
 484 (C) STRANDEDNESS: double  
 485 (D) TOPOLOGY: unknown  
 487 (ii) MOLECULE TYPE: DNA (genomic)  
 492 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
 E--> 494 GCCGCTGGAT GATCAGATTG CTGCTGCTAT GGTGTGAGG GGAATGGCTA AGGATGGGCA  
 W--> 495 60  
 E--> 497 GTTTGCTTTG AAGGATGATG CTGCTAAGGA TGGAGATAAA ACGGGGGTTG CTGCGGATGT  
 W--> 498 120  
 E--> 500 GAAAAATCCGA TTGACGCGGC TATTGGGGGT GCGGATGCTG ATGCTGCGGC GTTTAATAAG  
 W--> 501 180  
 E--> 503 GAGGGGATGA AGAAGGATGA TCAGATTGCT GCTGCTATGG TTCTGAGGGG AATGGCTAAG  
 W--> 504 240  
 E--> 506 GATGGGCAGT TTGCTTTGAC GAATAATGCT GC  
 E--> 507 272  
 509 (2) INFORMATION FOR SEQ ID NO: 12:  
 511 (i) SEQUENCE CHARACTERISTICS:  
 512 (A) LENGTH: 289 base pairs  
 513 (B) TYPE: nucleic acid

*same*  
↓

RAW SEQUENCE LISTING                      DATE: 05/12/2000  
 PATENT APPLICATION: US/09/445,803        TIME: 11:04:23

Input Set : A:\sequence.lst.txt  
 Output Set: N:\CRF3\05122000\I445803.raw

514            (C) STRANDEDNESS: double  
 515            (D) TOPOLOGY: unknown  
 517            (ii) MOLECULE TYPE: DNA (genomic)  
 522            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
 E--> 524 ACTGTTAAGA ATGCTGTTGA TATAATAAAG GCTGCTGCGG AAGCTGCAAG TGCTGCAAGT  
 W--> 525 60  
 E--> 527 GCTGCTACTG GTAGTGCAAG AATTGGGGAT GTTGTTAATG GTAATGGAGC AACAGCAAAA  
 W--> 528 120  
 E--> 530 GGTGGTGATG CGAAGAGTGT TAATGGGATT GCTAAGGGGA TAAAGGGGAT TGTGTGCTGCT  
 W--> 531 180  
 E--> 533 GCTGAGAAGG CTGATGCGAA GGAAGGGAAG TTGGATGTGG CTGGTGATGC TGGTGAAACT  
 W--> 534 240  
 E--> 536 AACAAAGGATG CTGGGAAGTT GTTTGTGAAG AACAAATGGTA ATGAGGGTA  
 E--> 537 289  
 539 (2) INFORMATION FOR SEQ ID NO: 13:  
 541            (i) SEQUENCE CHARACTERISTICS:  
 542                (A) LENGTH: 142 base pairs  
 543                (B) TYPE: nucleic acid  
 544                (C) STRANDEDNESS: double  
 545                (D) TOPOLOGY: unknown  
 547            (ii) MOLECULE TYPE: cDNA  
 550            (ix) FEATURE:  
 551                (A) NAME/KEY: CDS  
 552                (B) LOCATION: 2..142  
 554            (ix) FEATURE:  
 555                (A) NAME/KEY: mat\_peptide  
 556                (B) LOCATION: 2  
 559            (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
 E--> 561 G CCG CTT ACA AAT CCG ATT GAC GCG GCT ATT GGG GGG AGT GCG GAT  
 W--> 562 46  
 563    Pro Leu Thr Asn Pro Ile Asp Ala Ala Ile Gly Gly Ser Ala Asp  
 W--> 564        1                                5                                10                                15  
 W--> 566 CGT AAT GCT GAG GCG TTT GAT AAG ATG AAG AAG GAT GAT CAG ATT GCT  
 W--> 567 94  
 568 Arg Asn Ala Glu Ala Phe Asp Lys Met Lys Lys Asp Asp Gln Ile Ala  
 W--> 569                                20                                25                                30  
 W--> 571 GCT GCT ATG GTT CTG AGG GGA ATG GCT AAG GAT GGG CAG TTT GCT TTG  
 W--> 572 142  
 573 Ala Ala Met Val Leu Arg Gly Met Ala Lys Asp Gly Gln Phe Ala Leu  
 E--> 574                                35                                40                                45

*same*  
 ↓

## VERIFICATION SUMMARY

DATE: 05/12/2000

PATENT APPLICATION: US/09/445,803

TIME: 11:04:24

Input Set : A:\sequence.lst.txt

Output Set: N:\CRF3\05122000\I445803.raw

L:66 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:67 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:69 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:71 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:72 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:74 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:76 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:77 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:79 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:81 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:82 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:84 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:86 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:87 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:89 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:91 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:92 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:94 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:96 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:97 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:99 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:101 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:102 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:104 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:106 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:107 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:109 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:111 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:112 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:114 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:116 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:117 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:119 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:121 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:122 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:124 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:126 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:127 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:129 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:131 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:132 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:134 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:136 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:137 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:139 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:141 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
L:142 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:144 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1

## VERIFICATION SUMMARY

DATE: 05/12/2000

PATENT APPLICATION: US/09/445,803

TIME: 11:04:24

Input Set : A:\sequence.lst.txt

Output Set: N:\CRF3\05122000\I445803.raw

L:146 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
 L:147 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:149 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:151 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
 L:152 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:154 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:156 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
 L:157 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:159 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:161 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
 L:162 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:164 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:166 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:16  
 L:167 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:169 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:171 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:13  
 L:172 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:174 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:174 M:204 E: (24) Calc# of Bases differs from actual, LENGTH:Input:1047 Counted:0  
 L:269 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
 L:269 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
 L:270 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
 L:272 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
 L:272 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
 L:273 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
 L:275 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
 L:275 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
 L:276 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
 L:278 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
 L:278 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
 L:279 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
 L:281 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:5  
 L:281 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
 L:282 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
 L:282 M:204 E: (24) Calc# of Bases differs from actual, LENGTH:Input:283 Counted:0  
 L:299 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
 L:299 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
 L:300 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4  
 L:302 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
 L:302 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
 L:305 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
 L:305 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
 L:308 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
 L:308 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
 L:309 M:204 E: (24) Calc# of Bases differs from actual, LENGTH:Input:233 Counted:0  
 L:326 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
 L:326 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
 L:329 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
 L:329 M:333 E: Wrong sequence grouping, Amino acids not in groups!

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/445,803

DATE: 05/12/2000

TIME: 11:04:24

Input Set : A:\sequence.lst.txt

Output Set: N:\CRF3\05122000\I445803.raw

L:332 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
L:332 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:335 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:2  
L:335 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:336 M:204 E: (24) Calc# of Bases differs from actual, LENGTH:Input:194 Counted:0  
L:353 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
L:353 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:356 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
L:356 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:359 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
L:359 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:362 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
L:362 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:365 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
L:365 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:368 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
L:368 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:371 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:1  
L:371 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:372 M:204 E: (24) Calc# of Bases differs from actual, LENGTH:Input:369 Counted:0  
L:389 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
L:389 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:392 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
L:392 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:395 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:3  
L:395 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:396 M:204 E: (24) Calc# of Bases differs from actual, LENGTH:Input:142 Counted:0  
L:413 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
L:413 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:416 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
L:416 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:419 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
L:419 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:422 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:3  
L:422 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:423 M:204 E: (24) Calc# of Bases differs from actual, LENGTH:Input:210 Counted:0  
L:440 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:6  
L:440 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:443 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:446 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:449 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:450 M:204 E: (24) Calc# of Bases differs from actual, LENGTH:Input:236 Counted:0  
L:467 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:470 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:473 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:476 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:477 M:204 E: (24) Calc# of Bases differs from actual, LENGTH:Input:199 Counted:0  
L:494 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:497 M:333 E: Wrong sequence grouping, Amino acids not in groups!



## VERIFICATION SUMMARY

DATE: 05/12/2000

PATENT APPLICATION: US/09/445,803

TIME: 11:04:24

Input Set : A:\sequence.lst.txt

Output Set: N:\CRF3\05122000\I445803.raw

L:500 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:503 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:506 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:507 M:204 E: (24) Calc# of Bases differs from actual, LENGTH:Input:272 Counted:0  
L:524 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:527 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:530 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:533 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:536 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:537 M:204 E: (24) Calc# of Bases differs from actual, LENGTH:Input:289 Counted:0  
L:561 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:574 M:204 E: (24) Calc# of Bases differs from actual, LENGTH:Input:142 Counted:0